Stereotypes and Language Attitudes: Examining the Perception of German as "Harsh"

Undergraduate Research Thesis

Presented in Partial Fulfillment of the Requirements for Graduation "with Honors Research Distinction in German" in the Undergraduate Colleges of The Ohio State University

by

Megan Wadas

The Ohio State University

May 2020

Project Advisor: Dr. Carmen Taleghani-Nikazm, Department of Germanic Languages and Literatures

Abstract

Frequently described as guttural, harsh, or phlegmy, German has often been perceived negatively by native speakers of languages that do not have the same phonetic profile, such as English (Copy Cat Channel, 2013; Cronin, 2013). Indeed, even Charles V, the Holy Roman Emperor from 1519-1566, has been reported as remarking, "I speak Spanish to God, Italian to women, French to men, and German to my horse" (Jenkin, 2014). What does this phenomenon mean for the perception of not only the German language, but for the perception of German speakers and their cultures as a whole? Do preexisting cultural stereotypes of Germans sensitize laypeople to the perception of certain German sounds as "harsh", or does the language itself somehow bring these stereotypes to mind?

This study examined these questions using an online survey wherein labeled and unlabeled speech samples of German and Dutch were presented to various groups of respondents for an aesthetic rating either with or without priming of cultural stereotypes. It is hypothesized that participants will rate whatever language is labelled "German" more harshly and negatively aesthetically, no matter if the language was German or Dutch. Participants who rated the labeled language without stereotype prompting were hypothesized to rate the language less harshly. In addition, it was expected that participants would make more neutral judgements about the stereotypes that they were prompted to rate. The final hypothesis is the refutation of the "inherent value hypothesis", or the idea that some languages are inherently more logical, more beautiful, or more suited to certain applications than other languages (Ryan, Giles, & Hewstone, 1988). The final hypothesis is that the unlabeled languages will not generate the same ratings of aesthetic quality or stereotypical beliefs as when the language being spoken is presented in a labeled context.

Although the majority of these hypotheses were not supported by the data and that it appears that individual speaker differences made most of the difference in aesthetic and stereotype ratings, some interesting findings did come out of this study. Germans were consistently rated more “German” than Dutch people in every comparison, there do seem to be consistently salient “German” stereotypes, and there is some evidence to support the hypothesis that the label “German” does have some effect on how a language is perceived. This leads to the conclusion that the generation of aesthetic ratings of languages is a complex process that merits further, more in-depth study.

Stereotypes and Language Attitudes: Examining the Perception of German as "Harsh"

Literature Review

Language attitude study is a subfield of sociolinguistics that deals specifically with how people generate certain beliefs and attitudes towards speakers of a language or the culture a language is associated with. The complex process of how language is generated and interpreted between a speaker and a listener can be affected by countless factors and as a result, language attitude study is an extremely interdisciplinary field that incorporates research from social psychology, sociology, discourse analysis, communication, and anthropological linguistics (Cargile, Giles, Ryan, & Bradac, 1994). Some of the factors that are hypothesized to shape linguistic attitudes include extra-linguistic phenomena (such as gestures and body language), the hearer's behavior and emotional state, interpersonal history between the interlocutors, and, most importantly for this study, the language itself and the perceived cultural factors at play between the participants. Cargile et al. (1994) describe these factors in a "Process model of language attitudes" (Figure 1).

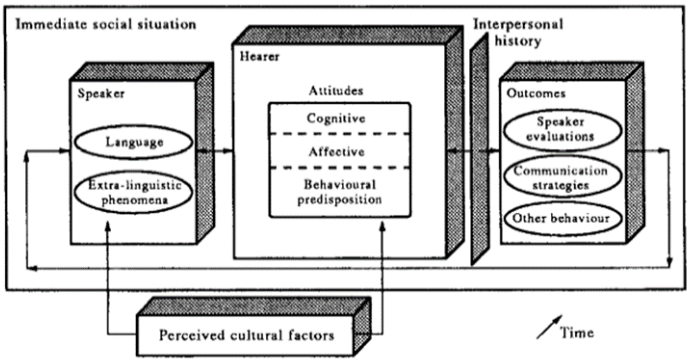

Figure 1. A Social Process Model of Language Attitudes

As is evident from this diagram, the pathways on which these various factors operate is bidirectional, meaning that when speakers produce language, listeners are instantaneously and selectively processing the various factors above in order to comprehend and respond to the speaker. Crucially, Cargile et al. (1994) posits that "speaker language does not inevitably trigger certain attitudes within the hearer, but rather hearers are actively involved in the process of selecting and attending to those language behaviors that meet their needs". Using this model, this study hypothesizes that when a native English speaker with no knowledge of German hears German and is aware of the language's identity, they selectively attend to the language behaviors that meet their needs: namely, the potentially subconscious need to internally confirm that common stereotypes about the sound of German being harsh are true. This study uses the above model as a pathway to investigate the effects of altering the language involved in a dialogue at hand in conjunction with manipulating the perceived cultural factors salient in a listener's mind. .

The notion that some amount of stereotype influence is present in these negative judgements of German is further supported by three ideas: that a) the "inherent value hypothesis" is dubiously sound at best, b) that the "imposed norm hypothesis" is better supported by research, and c) that a speaker's perceived ethnicity has been found to have a profound effect on how a language stimulus is interpreted (Giles, Bourhis, & Davies, 1979). The "inherent value hypothesis" is the idea that some languages possess inherent beauty, logic, or ability to express ideas. It has been shown by Giles et al. (p. 591, 1979), however, that inherent qualities of languages or universal perceptions of what makes a language beautiful may not exist: if they did, then "it may be that standard accents across cultures could have phonological features in common". However, the evidence that has been presented suggests that judgements of beauty associated with a language are instead based on some other notion of what a specific group of

people consider to be pleasing speech (Giles, Bourhis, & Taylor, 1977). In many cultures or groups, the language or dialect commonly labeled as more aesthetically pleasing also happens to be produced by the speakers in a culture that are accorded the highest status, whether that be economically, socially, or occupationally. This suggests that rather than there being some inherent beauty in a specific language or dialect, that a language or dialect attains a high prestige/aesthetically pleasing status through the "class or group which possesses it" (Giles, Bourhis, & Davies, 1979). The interactions between the sociocultural status of the speakers of a language affecting the perception of the language or dialect itself are known as the "imposed norms" that are present in a sociolinguistic landscape (Giles, Bourhis, & Davies, 1979)

When this group-specific intracultural and sociocultural context (the "imposed norms") surrounding various accents and dialects of a language disappear, so do the different relative value judgements associated with them (Trudgill & Giles, 1978). Trudgill & Giles (1978) propose that "cultural outsiders", that is, people who were not raised in nor belong to the specific people group under analysis, are not aware of the social connotations associated with different varieties of a language. Evidence for this claim was found in their study: "The fact that [participants] rated all the varieties [of a language] as approximately equal on aesthetic grounds [] strongly suggests that there are good reasons for arguing that all dialects are equally pleasant" (p. 180). Although all dialects of the same language may be judged as equally pleasant by the "cultural outsiders", the "cultural outsiders" still belong to an outgroup that has complex attitudes about the language in question, just not the attitudes that are formed within the society or culture that speaks said language. The imposed norm hypothesis stands in direct opposition to the inherent value hypothesis in that it suggests that rather than there being some kind of inherently beautiful or ugly qualities of a language, any judgements a listener makes about a

language are made as a result of social conditioning based on the "social connotations" of the group the speaker is perceived to belong to (Halliday, McIntosh, & Strevens, 1964).

The key word here is "perceived": the powerful ideas and stereotypes associated with a specific culture or language have been shown to shape judgements and attitudes towards the language itself and even factor into the ethnolinguistic vitality of one language or dialect over another (Giles, Bourhis, & Taylor, 1977). When discussing German culture and language with laypeople, media depictions or historical narratives about World War II or the Nazi Party are inevitably brought up and used as a symbol of German aggression and militarism (rewboss, 2014). This lasting concept of conflict between Third Reich Germany and the United States during WWII persists and could affect how the German language is perceived in the United States to this day given that there is evidence to show that conflict changes how people relate to their social or ethnic groups along linguistic lines (Kramer, 1992).

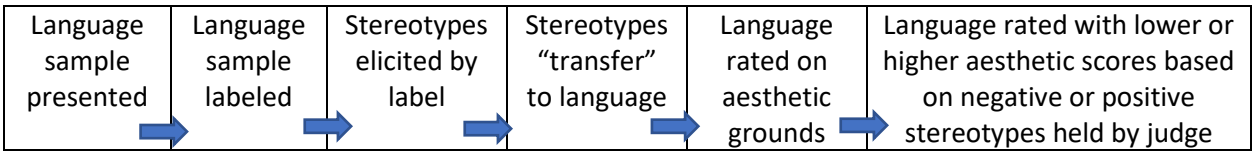

Figure 2. A hypothesized series of stages between the presentation of a language and the effect of the elicited stereotypes on the aesthetic rating

Method

The data for this research project was collected from a large pool of subjects who were asked to participate in an online survey. Participants were drawn from The Ohio State University population at large via flyers in academic buildings, an advertisement in a newsletter distributed by The Center for Languages, Literatures, and Cultures, and from the Linguistics Outside of the Classroom student research pool. These populations were chosen because they represent a diverse cross section of The Ohio State student body, especially given that many students from a variety of disciplines are completing their world language General Education requirement. Participants who are 18 or older who have English as their native language and no familiarity with German or Dutch were selected.

Using an online survey, participants evaluated recordings of German and Dutch. Dutch is a Germanic language with a similar phonetic inventory to German and hence was chosen as the control language for this experiment. The recordings used in this experiment were produced using the following speakers (Figure 3).

Although the German and Dutch speakers come from different variety backgrounds within their respective languages, the differences in accent and dialect should not make a difference in this experiment. A study by Giles et al. (1974) on the aesthetic rating of Greek by British adults with no knowledge of Greek investigated this issue, and found that while Greek speakers gave higher aesthetic ratings to the stereotypically “intelligent and sophisticated” Athenian variety of Greek (as compared to the Cretan variety), both were rated as equally pleasant by the British participants with no knowledge of Greek. In a similar study, Giles et al. (1974) showed the same results with non-French speaking British participants who rated various French varieties equally aesthetically. This led Giles et al. to describe the ideas of the “imposed

norm" hypothesis, described above, where the notion of a linguistically marked sociocultural status ceases to play a role in the aesthetic ratings of Greek or French by the British participants, who had no knowledge of the status differences that the different varieties of Greek or French might entail. This makes "cultural outsiders" the best and most unbiased judges of a language, even when different dialects of the same language are used for analysis. Therefore, native English-speaking judges with no knowledge of German or Dutch should not be affected by the different varieties of German and Dutch used by the native speakers in this study.

Figure 3. Speaker Information

Language	German		Dutch	
Gender	Male	Female	Male	Female
Age	25	24	50	51
Nationality	Austrian	German	Dutch	Dutch
Background	Born in Austria, native Austrian German speaker. Currently a doctoral student, has lived in the United States since 2019.	Born in Germany, native German speaker. Currently a master's student, has lived in the United States since 2019.	Born in the Netherlands, native Dutch speaker. Has lived in the United States since 1994 after graduating college in the Netherlands.	Born in the Netherlands, native Dutch speaker.

The participants were randomly divided into three experimental groups, each with two subgroups for the male and female recordings:

Group 1

A. Stereotype Priming: Subjects assessed their agreement or disagreement with stereotypes about the German culture on a Likert scale.

B. Language Stimulus: Subjects listened to a recorded language stimulus labeled as German. The stimulus was a sample of German produced by a native speaker of German.

C. Aesthetic Rating: The participants then rated the aesthetic qualities of the stimulus they were presented using a scored Likert scale.

This process was then repeated to have participants evaluate the other Germanic language, Dutch. Subjects were primed to think of their stereotypical beliefs about Dutch culture using the same German Stereotype Agreement Likert scale and rate a labeled sample of Dutch. One subgroup will rate samples of German and Dutch produced by the male native speakers and the other subgroup will rate samples of German and Dutch produced by the female native speakers.

Group 2

Participants in this group followed the same series of steps as listed above, however, with one key difference, that is, that the labels applied to the language stimuli were switched.

A. Stereotype Priming: Subjects assessed their agreement or disagreement with stereotypes about the German culture on a scored Likert scale.

B. Language Stimulus: Subjects listened to a recorded language stimulus labeled as German. The stimulus, however, was actually a sample of Dutch produced by a native speaker of that language.

C. Aesthetic Rating: The participants then rated the aesthetic qualities of the stimulus they were presented, also using a scored Likert scale.

This process was done in order to have participants evaluate the opposite condition: being primed with stereotypes about Dutch culture, then hearing a language stimulus of German labeled as Dutch. One subgroup rated samples of German and Dutch produced by the male native speakers and the other subgroup rated samples of German and Dutch produced by the female native speakers.

Group 3

Participants in this group were presented a language stimulus without any label and asked to rate it aesthetically as well as provide a rating of some supposed attributes of users of the "mystery" language. This was done for both German and Dutch and assessed using scales that include measures of stereotypes commonly found to be applied to German culture. The participants were also asked to guess what languages they thought the "mystery" languages were. Finally, the participants rated a language stimulus of German labeled as German for aesthetic attributes before providing information on their own personal stereotypical beliefs about German culture. This is in contrast to the first two groups, who were primed to be thinking stereotypically about each language's culture/people before rating the language sample aesthetically. One subgroup rated unlabeled samples of German and Dutch produced by the male native speakers followed by a labeled sample of German produced by the female German speaker, while the other subgroup rated unlabeled samples of German and Dutch produced by the female native speakers followed by a labeled sample of German produced by the male German speaker.

(Two subgroups were made out of each of these three groups, one that rated the male speakers' language samples and one that rated the female speakers' language samples. For Group 3, two subgroups were also made, one for the male unlabeled language samples and the female

labeled German language sample, and one for the female unlabeled language samples and the male labeled German language sample.)

All groups provided information on the languages they speak and cultures they have familiarity with as well as basic demographic information such as age, year in college, major, and whether the participant has ever lived outside of the United States. This allowed for the participants to be characterized and described appropriately. The estimated average time to complete the survey was 15-20 minutes. The following table (Figure 4) provides a summary/an overview of the groups and phases of the survey.

Figure 4. Experimental Setup

Group 1 Phase 1			Group 1 Phase 2		
Step 1 GSA Rating/Stereotype Priming	Step 2 Language Sample (Labeled as German)	Step 3 Aesthetic Rating of	Step 1 GSA Rating/Stereotype Priming	Step 2 Language Sample (Labeled as Dutch)	Step 3 Aesthetic Rating of
German	German	German w/ GP	Dutch	Dutch	Dutch w/ DP
Group 2 Phase 1			**Group 2 Phase 2**		
Step 1 GSA Rating/ Stereotype Priming	Step 2 Language Sample (Labeled as German)	Step 3 Aesthetic Rating of	Step 1 GSA Rating/Stereotype Priming	Step 2 Language Sample (Labeled as Dutch)	Step 3 Aesthetic Rating of
German	Dutch	Dutch w/ GP	Dutch	German	German w/ DP

Group 3 Phase 1: German, Unlabeled/No Priming			Group 3 Phase 2: Dutch, Unlabeled/No Priming			Group 3 Phase 3: German, Labeled/No Priming		
Step 1 Language Sample	Step 2 Aesthetic Rating	Step 3 GSA Rating	Step 1 Language Sample	Step 2 Aesthetic Rating	Step 3 GSA Rating	Step 1 Language Sample	Step 2 Aesthetic Rating	Step 3 GSA Rating

Note: GSA refers to German Stereotype Agreement

Likert Scales

German Stereotype Agreement Scale

The German Stereotype Agreement (GSA) Likert scale asked participants to answer the following questions, depending on which experimental group the participant is in and whether or not the language stimulus is labeled or unlabeled, such as the conditions in Group 3:

a. *Please rate how you perceive Germans/German culture;*

b. *Please rate how you perceive Dutch people/Dutch culture; or*

c. *Please rate how you perceive the people who speak this language/the culture this language belongs to.*

In all cases, participants were asked to rate the language or culture being evaluated using a 10-point Likert scale. The higher the GSA score, the more "stereotypically German" the language and culture is percieved to be. (Figure 5).

The axes on which the stereotypes were rated are Militaristic vs. Peaceful, Polluting vs. Eco-Friendly, Strict vs. Permissive, Serious vs. Funny, Heavy vs. Low Alcohol Consumption, Stagnant vs. Innovative, Uneducated vs. Intelligent, Greedy vs. Generous, Unequal vs. Egalitarian, Logical vs. Emotional, Disorganized vs. Orderly, Poor vs. Rich, and Uncultured vs. Rich Cultural Traditions. These items were chosen to prime participants to be thinking about their own stereotypical beliefs about certain cultures overall, and were selected based on stereotypes reflected a survey that examined the stereotypes of college learners of German about Germans and Germany (Schulz & Haerle, 1995). However, the attributes that were scored as 10 points, reflecting more "stereotypical German-ness", were *Militaristic, Eco-Friendly, Strict,*

Serious, Heavy Alcohol Consumers, Innovative, Intelligent, Greedy, Egalitarian, Logical, Organized, Rich, Cultured, and Punctual.

	1	2	3	4	5	6	7	8	9	10	Don't know/No opinion
1 = Militaristic vs. 10 = Peaceful	○	○	○	○	○	○	○	○	○	○	○
1 = Polluting vs. 10 = Eco-friendly	○	○	○	○	○	○	○	○	○	○	○
1 = Strict vs. 10 = Permissive	○	○	○	○	○	○	○	○	○	○	○
1 = Serious vs. 10 = Funny	○	○	○	○	○	○	○	○	○	○	○
1 = Heavy alcohol consumers vs. 10 = Light alcohol consumers	○	○	○	○	○	○	○	○	○	○	○

	1	2	3	4	5	6	7	8	9	10	Don't know/No opinion
1 = Stagnant vs. 10 = Innovative	○	○	○	○	○	○	○	○	○	○	○
1 = Uneducated vs. 10 = Intelligent	○	○	○	○	○	○	○	○	○	○	○
1 = Greedy vs. 10 = Generous	○	○	○	○	○	○	○	○	○	○	○
1 = Unequal vs. 10 = Egalitarian	○	○	○	○	○	○	○	○	○	○	○
1 = Logical vs. 10 = Emotional	○	○	○	○	○	○	○	○	○	○	○

	1	2	3	4	5	6	7	8	9	10	Don't know/No opinion
1 = Disorganized vs 10 = Orderly	○	○	○	○	○	○	○	○	○	○	○
1 = Poor vs. 10 = Rich	○	○	○	○	○	○	○	○	○	○	○
1 = Uncultured vs. 10 = Cultured	○	○	○	○	○	○	○	○	○	○	○
1 = Late vs. 10 = Punctual	○	○	○	○	○	○	○	○	○	○	○

Figure 5. The German Stereotype Agreement (GSA) Likert scale, as it appears on the online survey (unscored)

Aesthetic Rating Scale

The aesthetic rating scale (AR), a 10-point Likert scale asked the following questions:

a) *Do you find this language (German/Dutch/unlabeled) to be…* Halting vs. Melodic, Ugly vs. Beautiful, Guttural vs. Flowing, Harsh vs. Gentle, Disorganized vs. Rhythmic, Annoying vs. Soothing, and Unpleasant vs. Pleasant.

The attributes with the more positive connotation are scored for 10 points, resulting in a higher score implying a "more attractive" aesthetic rating (Figure 6).

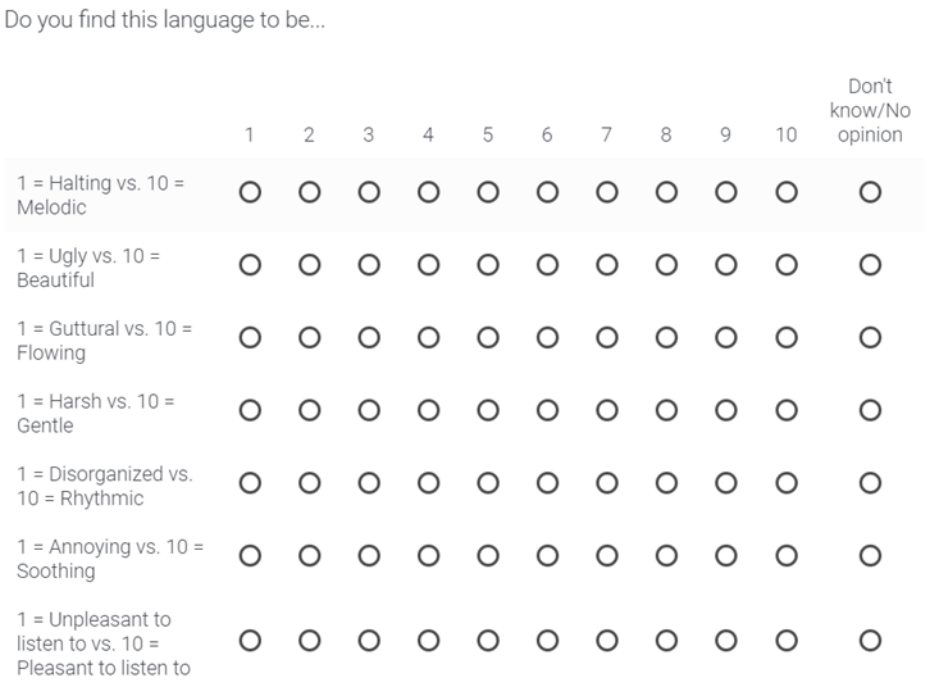

Figure 6. The Aesthetic Rating (AR) scales

Analysis

The data collected by the GSA (German Stereotype Agreement) and AR (Aesthetic Rating) Likert scales were analyzed using independent two-tailed t-tests and Pearson's Product-Moment Correlations across multiple comparisons (Figure 7). The color coding in Figure 6 is aligned with Figure 3 for the different groups and the different data sets.

First, the ARs of the different groups were compared against each other using independent two-tailed t-tests. By comparing the ARs of the various groups across the experimental conditions (when the language is correctly labeled, deliberately mislabeled, or unknown), conclusions can be drawn about what effect the label "German" and its associated cultural stereotypes have on the aesthetic perception of German or Dutch. This addresses the first hypothesis: that any language sample labeled as "German" will be rated more "harshly" and negatively aesthetically, regardless of whether if the language was German or Dutch. In addition, this comparison addressed part of the second hypothesis: that participants who rated the labeled language without stereotype prompting would rate the language to be less harsh. The bolded sections of Figure 6 display this process. The label "German" is examined in the first bolded section and the label "Dutch" in the second. The comparison with Dutch was done in order to determine if any differences in the AR are a result of Dutch being the control language rather than the label "German" having any significant effect on the AR.

Then, Pearson's Product-Moment Correlations were calculated between both individual stereotypical attributes (Militaristic, Eco-Friendly, Strict, Serious, Heavy Alcohol Consumers, etc.) and the GSA score and between the overall GSA score and the AR score. By making these comparisons, the stereotypes that are salient in listeners' minds while making these stereotype and aesthetic judgements can be determined. This is designed to address the other part of the

second hypothesis: that participants will make more neutral judgements about the stereotypes that they were prompted to rate when the language they are rating is unlabeled. Finally, it was hypothesized that unlabeled languages will not generate the same ratings of aesthetic quality or stereotypical beliefs as when the language being spoken is presented in a labeled context.

Figure 7. Analysis

Stereotypes Generated	vs.	Stereotypes Primed		Do the stereotypes about the German language and culture play a role in the aesthetic rating, or does the language itself influence the stereotypes?
German primed with GP[1]	vs.	German primed with DP[2]		Do the stereotypes about the German language influence the aesthetic rating of German?
German primed with GP		Dutch primed with GP		Consequently, do these two show similar aesthetic ratings based on the application of the label "German"?
Totally Unlabeled German	vs.	Stereotype Primed, Unlabeled German	Stereotype Unprimed, Labeled German	Does presenting German without a label result in differences in the aesthetic rating as compared to German rated with knowledge of the language's identity?
Dutch primed with DP	vs.	Dutch primed with GP		Do the stereotypes about the German language influence the aesthetic rating of Dutch?
Dutch primed with DP		German primed with DP		Consequently, do these two show similar aesthetic ratings based on the application of the label "Dutch"?
Inquiry into the Inherent Value Hypothesis				Data collected in the interest of determining whether or not the inherent qualities of German influence these aesthetic ratings

[1] GP- German stereotype rating prompt
[2] DP- Dutch stereotype rating prompt

Results

Data was collected from a total of 115 participants. Data from 13 non-native English speakers and 11 people who spoke German or Dutch were discarded, resulting in a participant number of n = 91. The majority of participants were of traditional college age.

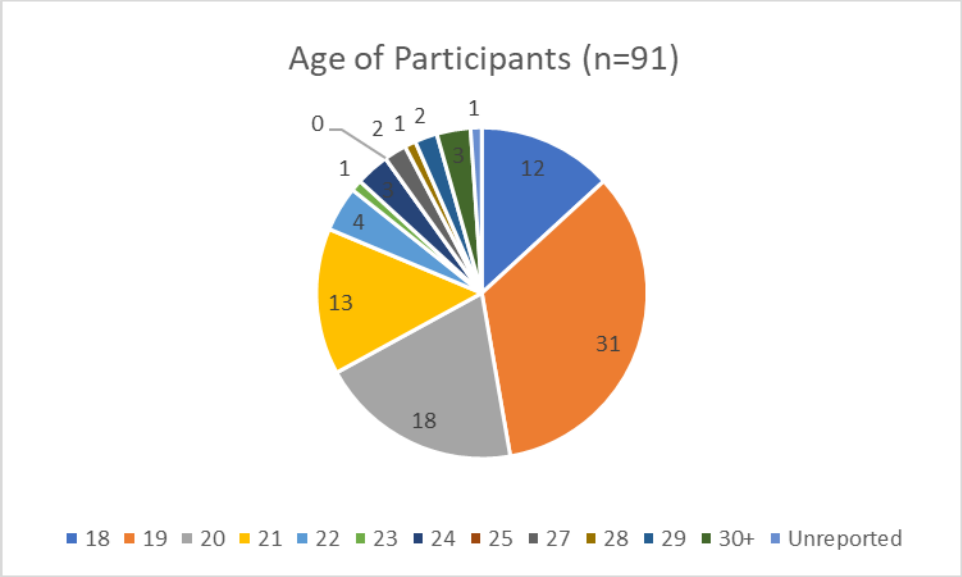

Figure 8. Ages of participants

In addition, the participants represented a variety of majors, including Linguistics, Speech and Hearing Sciences, Teaching English to speakers of other languages (TESOL), Neuroscience, and Education, among others.

The vast majority of participants reported to have studied multiple languages or cultures, the most common language studied being Spanish. A few praticipants (n = 10) lived abroad, suggesting a degree of openness to learning a new language or being immersed in a new culture.

Aesthetic Rating Comparisons

Male Speaker Subgroup AR Comparisons

The male speaker sample of German labeled as German (AR mean = 5.857) was rated to be significantly more aesthetically pleasing than the male speaker sample of Dutch labeled as Dutch (AR mean = 4.483, p = 4.44e-07). When the AR of German labeled as German (AR mean = 5.857) was compared to the same language sample in the unlabeled condition (AR mean = 6.485), the unlabeled sample was found to be significantly more aesthetically pleasing (p = 0.046). This was also the case when both languages were unlabeled, with unlabeled German (AR mean = 6.485) rated significantly more aesthetically pleasing than unlabeled Dutch (AR mean = 4.699, p = 1.63e-08). Dutch was also rated significantly more aesthetically pleasing when labeled as German (AR mean = 5.056) than correctly as Dutch (AR mean = 4.483, p = 0.033).

Gender	Comparison 1	Mean Aesthetic Rating	Comparison 2	Mean Aesthetic Rating	df	p-value
Male	German as German	5.857	Dutch as Dutch	4.483	233	4.44E-07
			German as Dutch	6.205	205	0.227
	German as German	5.857	German Unlabeled	6.485	218	0.046
			German w/o stereotype	5.923	221	0.826
	Dutch as Dutch	4.483	Dutch as German	5.056	203	0.033
		4.483	Dutch Unlabeled	4.699	217	0.389
	German Unlabeled	6.485	Dutch Unlabeled	4.699	202	1.63E-08

Figure 9. AR t-test comparisons, male speaker subgroups

Female Speaker Subgroup AR Comparisons

The female speaker sample of Dutch labeled as Dutch (mean = 6.982) was rated significantly more aesthetically pleasing than the female speaker sample of German labeled as German (AR mean =5.764, p = 5.54e-05). This was also the case when both languages were unlabeled, with unlabeled Dutch (AR mean = 7.210) rated to be significantly more aesthetically

pleasing than unlabeled German (AR mean = 5.875, p = 4.67e-06). No other AR comparisons returned significant results.

Gender	Comparison 1	Mean Aesthetic Rating	Comparison 2	Mean Aesthetic Rating	df	p-value
Female	German as German	5.764	Dutch as Dutch	6.982	218	5.54E-05
			German as Dutch	5.238	213	0.081
			German Unlabeled	5.875	212	0.735
			German w/o stereotype	5.660	211	0.756
	Dutch as Dutch	6.982	Dutch as German	6.800	213	0.467
			Dutch Unlabeled	7.210	213	0.359
	German Unlabeled	5.875	Dutch Unlabeled	7.210	207	4.67E-06

Figure 10. AR t-test comparisons, female speaker subgroups

Overall

There were no significant differences between the ARs across any of the different experimental groups. There was also no significant difference between the mean AR of the German male speaker sample (AR mean = 5.857) and the German female speaker language sample (AR mean = 5.764). However, the male Dutch speaker's language sample (mean = 4.483) was rated to be significantly less aesthetically pleasing than the female Dutch speaker's language sample (mean = 6.982, p = 2.20e-6).

Gender	Comparison 1	Mean Aesthetic Rating	Comparison 2	Mean Aesthetic Rating	df	p-value
Overall	German as German	5.813	Dutch as Dutch	5.699	453	0.588
			German as Dutch	5.679	420	0.524
			German Unlabeled	6.176	432	0.110
			German w/o stereotype	5.792	434	0.929
			Dutch as German	6.000	421	0.405
	Dutch As Dutch	5.699	Dutch as German	6.000	418	0.155
			Dutch Unlabeled	5.966	432	0.211
			German as Dutch	5.679	417	0.916
	German Unlabeled	6.176	Dutch Unlabeled	5.966	411	0.355
Mixed	German as German Male	5.857	German as German Fema	5.764	227	0.771
	Dutch as Dutch Male	4.483	Dutch as Dutch Female	6.982	224	2.20E-16

Figure 11. AR t-test comparisons, Overall (male and female) and Mixed Gender subgroups

German Stereotype Agreement Comparisons

Between Nationality/Language-Specific Prompts

In every comparison and every subgroup but one, the GSA mean was significantly higher for the prompt that asked participants to rate how they perceive Germans/German culture than for the prompt that asked participants to rate how they perceive Dutch people/Dutch culture. In addition, the participants in Group 1, who rated the female language samples, rated Dutch people (GSA mean = 6.349) significantly more "German" than the participants in Group 1 who rated the language samples for the male speakers (GSA mean = 5.728, p = 0.017). The only group in which this difference was not seen was between the Group 1 participants who rated the male speaker samples (GSA mean = 7.000) and the Group 1 participants who rated the female speaker samples (GSA mean = 6.868, p = 0.530).

	German- Group 1M	Dutch- Group 1M	Dutch- Group 1F	Dutch- Group 1M
GSA Mean	7.000	5.728	5.728	6.349
	German- Group 1F	Dutch- Group 1F	German- Group 1F	German- Group 1M
GSA Mean	6.868	6.349	6.868	7.000
p-value	0.530	0.017	3.44E-07	0.0066723
t-value	0.629	-2.407	-5.193	-2.728
df	397	342	364	375
GSA Mean	Dutch- Group 1 Overall	Dutch- Group 2 Overall	Dutch- Groups 1 & 2 Combined	
	6.043	5.931	5.991	
GSA Mean	German- Group 1 Overall	German- Group 2 Overall	German- Groups 1 & 2 Combined	
	6.932	6.80	6.872	
p-value	5.96E-08	1.02E-06	2.61E-13	
t-value	-5.476	4.935	7.386	
df	741	636	1379	

Figure 12. t-test comparisons between GSA Means

Between Nationality/Language-Specific Prompts and Unlabeled Prompts

There was only one significant difference in GSA means between the prompts that asked participants to rate how they perceive Germans/German culture and the prompts that asked participants to rate how they perceive "the people who speak this language/the culture this language belongs to": between the Group 1M participants (GSA mean = 7.000) and Group 3M participants, who were prompted to rate the unlabeled male sample of German stereotypically (GSA mean = 6.474, p = 0.017).

	German- Group 1F	German- Group 1M	German- Group 1 Overall
GSA Mean	6.868	7.000	6.932
	German- Group 3F	German- Group 3M	German- Group 3 Overall
GSA Mean	6.978	6.474	6.733
p-value	0.625	0.017	0.206
t-value	-0.489	2.392	1.266
df	380	375	757

Figure 13. t-tests between GSA means between labeled and unlabeled samples

Between GSA Means Rated Before or After Listening to the Language Samples

No significant differences were found in GSA means between the groups that rated German stereotype agreement before listening to the labeled German language sample and the groups that listened to the labeled German language sample first before rating their agreement with German stereotypes.

Individual Stereotype Comparisons

Group 1M

For the participants that rated the male language sample of German labeled as German, the stereotypes that were found to be significantly positively correlated with the overall GSA mean were Strict (r = 0.686, p = 0.005), Serious (r = 0.544, p = 0.036), Organized (r =0.817, p = 1.11e -04), and Punctual (r = 0.745, p =0.004). This means that these stereotypes were the ones that had the most bearing on what "Germanness" meant to the participants.

Innovative (r = 0.605, p = 0.028), Organized (r = 0.614, p = 0.011), and Punctual (r = 0.787, p = 0.004) were found to be significantly positively correlated with the AR.

German as German, Group 1M (Individual Stereotype Correlation Coefficients (r))					
	Comparison Mean			Comparison Mean	
Stereotype	GSA	AR	Stereotype	GSA	AR
Militaristic	0.28	-0.02	Greedy	0.20	-0.03
Eco-Friendly	0.42	0.26	Egalitarian	-0.36	0.04
Strict	0.69**	0.11	Logical	0.33	-0.24
Serious	0.54*	-0.06	Organized	0.82***	0.61*
Heavy Alcohol Consumers	0.50	0.14	Rich	0.37	0.15
Innovative	0.42	0.61*	Cultured	0.47	0.20
Intelligent	0.19	0.29	Punctual	0.74**	0.79**

*p < 0.05, **p < 0.01, *** p < 0.001

Figure 14. German as German, Group 1M Individual Stereotype Correlations

Group 1F

The group that rated the female language sample of German labeled as German had markedly different results. The stereotypes that were significantly positively correlated with the GSA were Strict ($r = 0.515$, $p = 0.050$), Serious ($r = 0.620$, $p = 0.010$), Heavy Alcohol Consumers ($r = 0.671$, $p = 0.006$), and Punctual ($r = 0.683$, $p = 0.014$). These stereotypes are very similar to the ones found to be significant for the group that rated male speaker, however, the stereotypes that ended up being correlated with the aesthetic rating were very different.

The stereotypes that were significantly correlated with the AR were Militarism ($r = -0.597$, $p = 0.019$), Strict ($r = -0.542$, $p = 0.037$), Intelligent ($r = 0.571$, $p = 0.033$), and Egalitarian ($r = 0.569$, $p = 0.034$). Negative stereotypes like being Militaristic and Strict that were not found to affect the AR rating of the male speaker sample were found to be negatively correlated with theh AR for the female speaker sample, implying that gender differences affected the AR as well.

German as German, Group 1F (Individual Stereotype Correlation Coefficients (r))					
	Comparison Mean			Comparison Mean	
Stereotype	GSA	AR	Stereotype	GSA	AR
Militaristic	0.10	-0.60*	Greedy	0.71**	-0.55
Eco-Friendly	0.55	0.65*	Egalitarian	-0.24	0.57*
Strict	0.51*	-0.54*	Logical	0.54*	-0.38
Serious	0.62**	-0.44	Organized	0.41	-0.05
Heavy Alcohol Consumers	0.67**	-0.48	Rich	0.58*	0.28
Innovative	0.35	0.44	Cultured	-0.07	0.42
Intelligent	0.06	0.57*	Punctual	0.68*	0.01

*p < 0.05, **p < 0.01, *** p < 0.001

Figure 15. German as German, Group 1F Individual Stereotype Correlations

Group 1 Overall

For the Group 1 participants overall, Eco-Friendly (r = 0.428, p = 0.037), Strict (r = 0.589, p = 0.001), Serious (r = 0.495, p = 0.005), Heavy Alcohol Consumers (r = 0.531, p = 0.002), Greedy (r = 0.431, p = 0.025), Logical (r = 0.451, p = 0.014), Organized (r = 0.717, p = 8.31e-06), Rich (r = 0.411, p = 0.024), and Punctual (r = 0.710, p = 1.49e-04) were all found to be significantly correlated with the GSA. This means that these stereotypes were the ones found to be most prevalent and salient for "Germaness" for this group.

Being Eco-Friendly (r = 0.445, p = 0.030) and Innovative (r = 0.503, p = 0.008) were found to be significantly correlated with the AR.

German as German, Group 1 Overall (Individual Stereotype Correlation Coefficients (r))					
	Comparison Mean			Comparison Mean	
Stereotype	GSA	AR	Stereotype	GSA	AR
Militaristic	0.17	-0.40*	Greedy	0.43*	-0.33
Eco-Friendly	0.43*	0.44*	Egalitarian	-0.28	0.37
Strict	0.59**	-0.22	Logical	0.45*	-0.30
Serious	0.49**	-0.20	Organized	0.72***	0.30
Heavy Alcohol Consumers	0.53**	-0.12	Rich	0.41*	0.21
Innovative	0.37	0.50**	Cultured	0.20	0.34
Intelligent	0.15	0.44*	Punctual	0.71***	0.36

*p < 0.05, **p < 0.01, *** p < 0.001

Figure 16. German as German, Group 1 Overall Individual Stereotype Correlations

Group 1 and 2 Combined

For the GSA data collected on German in both Groups 1 and 2, the majority of the individual stereotypes were shown to be correlated with the overall GSA mean. The stereotypes that were found to be significantly correlated with the overall GSA mean for German were identical to those found to be significantly correlated with the overall GSA mean for Dutch with two exceptions: Heavy Alcohol Consumers was significantly positively correlated with the overall GSA mean for German ($r = 0.55$, $p = 1.30e-05$) but not for Dutch, whereas Innovative was significantly positively correlated with the overall GSA mean for Dutch but not for German ($r = 0.51$, $p = 3.29e-04$).

Groups 1 and 2 Overall (Individual Stereotype vs. GSA Correlation Coefficients (r)) German			
Stereotype	GSA	Stereotype	GSA
Militaristic	0.20	Greedy	0.24
Eco-Friendly	0.32*	Egalitarian	-0.25
Strict	0.45**	Logical	0.48***
Serious	0.43**	Organized	0.64***
Heavy Alcohol Consumers	0.55***	Rich	0.52***
Innovative	-0.01	Cultured	0.31*
Intelligent	0.29	Punctual	0.67***

*p < 0.05, **p < 0.01, *** p < 0.001

Figure 17. German Stereotype/GSA Correlations

Groups 1 and 2 Overall (Individual Stereotype vs. GSA Correlation Coefficients (r)) Dutch			
Stereotype	GSA	Stereotype	GSA
Militaristic	0.20	Greedy	0.20
Eco-Friendly	0.27	Egalitarian	0.03
Strict	0.39**	Logical	0.48***
Serious	0.37*	Organized	0.58***
Heavy Alcohol Consumers	0.01	Rich	0.61***
Innovative	0.51***	Cultured	0.33*
Intelligent	0.61	Punctual	0.59**

*p < 0.05, **p < 0.01, *** p < 0.001

Figure 18. Dutch Stereotype/GSA Correlations

Group 3 Overall

The stereotypes found to be correlated with the GSA means for Group 3 participants, who rated the unlabeled sample of German aesthetically before giving their rating on the GSA scale on "the people who speak this language/the culture this language belongs to", were Eco-Friendly ($r = 0.457$, $p = 0.025$), Strict ($r = 0.558$, $p = 0.002$), Serious ($r = 0.476$, $p = 0.009$), Heavy Alcohol Consumers ($r = 0.623$, $p = 0.002$), Intelligent ($r = 0.523$, $p = 0.004$), Rich ($r = 0.484$, $p = 0.019$), Cultured ($r = 0.402$, $p = 0.038$), and Punctual ($r = 0.753$, $p = 3.37e\text{-}05$).

The stereotypes found to be correlated with the AR means were Militaristic ($r = -0.467$, $p = 0.016$), Heavy Alcohol Consumers ($r = -0.526$, $p = 0.012$), Intelligent ($r = 0.561$, $p = 0.002$), Greedy ($r = -0.467$, $p = 0.021$), and Cultured ($r = 0.466$, $p = 0.014$). In sum, people who rated Germans as more Militaristic, Heavy Alcohol Consumers, or Greedy rated German in turn as less aesthetically pleasing. However, these were not the same stereotypes found to be significantly correlated with the GSA mean overall, implying a disconnect between the salient German stereotypes and which of them affect the AR.

German Unlabeled, Group 3 Overall (Individual Stereotype Correlations (r))					
	Comparison Mean			Comparison Mean	
Stereotype	GSA	AR	Stereotype	GSA	AR
Militaristic	-0.23	-0.47*	Greedy	0.06	-0.47*
Eco-Friendly	0.46*	-0.03	Egalitarian	0.31	0.14
Strict	0.56*	-0.16	Logical	0.25	-0.03
Serious	0.48**	-0.12	Organized	0.32	0.23
Heavy Alcohol Consumers	0.62**	-0.53*	Rich	0.48*	0.05
Innovative	0.32	0.16	Cultured	0.40*	0.47*
Intelligent	0.52**	0.56**	Punctual	0.75***	0.09

*p < 0.05, **p < 0.01, *** p < 0.001

Figure 19. German Unlabeled, Group 3 Overall (Individual Stereotype Correlations (r))

Correlations with the Individual Aesthetic Qualities and the AR Mean

In Group 1M, only “Pleasant” was significantly positively correlated with the GSA mean ($r = 0.508$, $p = 0.037$). The GSA mean and AR means were not significantly correlated. In Group 1 Overall, only “Beautiful” was significantly positively correlated with the GSA mean ($r = 0.364$, $p = 0.040$).

In Groups 1F, 3M, 3F, and 3 Overall, there were no significant correlations between the GSA mean and any of the aesthetic qualities or between the GSA mean and the AR mean. This means that there does not appear to be a direct link between how stereotypically “German” people rate Germans and how positively or negatively they percieve their speech.

Discussion

It was hypothesized that any language sample labeled as German would be rated to be less aesthetically pleasing than the same language samples labeled as Dutch or the same language samples in an unlabeled condition. This was not found to be the case. No significant differences in AR were found between the experimental groups cross-linguistically or in the conditions with the switched labels, suggesting that the first hypothesis is unlikely.

Additionally, speakers that rated language samples of German aesthetically without stereotype prompting were not found to rate the samples aesthetically any more or less harshly than the group that rated the samples with stereotype prompting. The first part of second hypothesis, that the prompting of German stereotypes would result in a lower AR score and its reverse, that the lack of stereotype prompting in the unlabeled and switch order conditions would result in a higher AR score, was not supported by the data.

Oddly enough, however, in the t-test comparison between the German GSA means for Group 1M (correctly labeled male speaker samples) and Group 3M (unlabeled male speaker samples), the GSA mean for Group 3M (GSA mean = 6.474) was significantly lower than Group 1M (GSA mean = 7.000, $p = 0.017$), implying that there may be some merit to the second part of the second hypothesis: that the unlabeled samples of German would result in less stereotypical ratings than in the labeled condition. This result, however, was not seen in the female speaker groups or in the combined male and female speaker group overall.

Nevertheless, the AR independent two-tailed t-tests suggest that individual speaker differences made most of the differences in AR rather than the label “German” having any particular effect. The fact that the male and female German speakers’ samples were not found to

have significantly different ARs from each other while the Dutch speakers' did implies that something about the male Dutch speaker's voice was perceived to be particularly aesthetically unappealing while the female Dutch speaker's voice was perceived to be particularly aesthetically pleasing.

The effects of the Dutch voices being outliers was also seen in the male and female subgroups. In the male speaker subgroup, the German voice was consistently rated more aesthetically appealing than the Dutch voice when the two were compared against each other in both the correctly labeled and unlabeled conditions. In the female speaker subgroup, the same but opposite pattern occurred: the Dutch voice was consistently rated more aesthetically appealing than the German voice when the two were compared against each other in both the correctly labeled and unlabeled conditions.

Evidence for the idea that the label "German" does have some effect on the AR does still exist, though: When the AR of the male speaker sample of German labeled as German (AR mean = 5.857) was compared to the AR of experimental group that rated the same sample in the unlabeled condition, the unlabeled sample was rated to be significantly more aesthetically pleasing (AR mean = 6.485, $p = 0.046$). On the other hand, the male speaker sample of Dutch labeled as German (AR mean = 5.056) was rated significantly to be more aesthetically pleasing than the same sample labeled as Dutch (AR mean = 4.483, $p = 0.033$). This is contrary to the hypothesis that any language sample rated as "German" would result in a lower AR, but seems to make sense in the context of the male Dutch voice being an outlier in terms of low AR score.

Part of this discrepancy in hypothesized AR scores might be due to greater familiarity with German than with Dutch amongst the participants. In Group 3, where participants were asked to rate the unlabeled language samples, of the participants overall (n= 30), the vast

majority of those who made a guess at all (n=20) were able to correctly identify the unlabeled sample of German as German (n = 17). Incorrect guesses included Italian (n =1) and Dutch (n = 2). Most of the participants who correctly guessed German gave reasons: “I have German friends and family”, “I have heard the sounds in movies and media”, “words were recognized”, “das, wunderbar” (German words for *the* and *wonderful*), “I have a friend who speaks German”, “we have hosted German exchange students”, “seemed very halting but somehow familiar”, and “I grew up with an Oma and Opa who had a similar accent” (*Oma* and *Opa* being the German words for grandmother and grandfather).

However, only 11 participants made a guess on the Dutch samples and only 2 guessed correctly that the second language sample was Dutch. Incorrect guesses included Scots (n = 1), “Scottish or Irish Gaelic” (n = 1), Danish (n = 2), Norwegian (n =1), “Slavic of some kind” (n = 1), Yiddish (n = 1), “A European language, maybe German” (n = 1), and “German again” (n = 1).

This suggests that whatever “German” sounds like to laypeople has been well enough established to the point that without any prompting about the language’s identity, people with no formal instruction in German are able to correctly guess the language. This may also explain why the GSA correlations for the Group 3 unlabeled samples so closely matched the Group 1 labeled GSA correlation data: the participants simply guessed that the language being spoken was German and applied stereotypes in a very similar manner to the participants who rated German in the labeled condition. The individual stereotypes that were significantly correlated with the GSA Mean in both Group 1 (labeled) and Group 3 (unlabeled) were Eco-Friendly, Strict, Serious, Heavy Alcohol Consumers, Rich, and Punctual. The individual stereotypes that were

significantly correlated with the AR in both groups were Militaristic and Intelligent, which were negatively and positively correlated with the AR respectively.

What is perhaps most important to note is that Germans had a significantly higher GSA mean overall than Dutch people in every possible comparison other than between the male and female German speaker sample groups themselves. Participants had clear ideas about Germans that resulted in GSA scores that were shown by a t-test to be significantly higher than their ratings of Dutch people, reinforcing the idea that participants do perceive the two cultures differently. Additionally, there are stereotypes across the groups that are significantly correlated the GSA mean when analyzed both individually and as a larger group, suggesting that there are stereotypes about Germans that are consistently salient in the population's minds.

When the German GSA data for Groups 1 and 2 was assessed together, there were moderately strong correlations for Eco-Friendly, Strict, Serious, Heavy Alcohol Consumers, Logical, Organized, Rich, Cultured, and Punctual, which is a very similar set to the list above, though not identical. Militaristic, the stereotype that was the original inspiration for the hypothesis that laypeople would find German to be harsh given the culture's history in the 20th century, was not significantly correlated with the GSA mean in this grouping.

Although there are stereotypes that appear to be consistently salient in people's minds when they are prompted to think about Germans or guess whether an unlabeled language is German, these stereotypes do not necessarily directly affect the AR in the manner hypothesized. In fact, in no case was the overall GSA Mean significantly correlated with the AR Mean in any group, either positively or negatively. However, there were certain negative stereotypes (Militaristic and Strict) that were significantly negatively correlated with the AR, but only for the

female German speaker, implying that the gender of the speaker may play a bigger role in how speech is perceived than the identity of the language.

Therefore, it can be inferred that the way people perceive language is the result of other, more complex interactions between speakers and hearers and their cultural upbringings rather than simply expanding positive and negative stereotypes about a group of people to encompass the perception of the language as well.

Limitations

This study did have some notable limitations. A factor that may affect the data is the population sample. Although the study was available to the entire Ohio State student population, the overwhelming majority of the participants were recruited from the Linguistics Outside of the Classroom student research pool as well as from a Speech and Hearing Sciences extra credit opportunity. The participants were predominantly of traditional college age and had more often than not studied more than one world language. Many were Linguistics majors or minors as well, further predisposing the sample population to be hesitant to label any culture stereotypically or any language as particularly unpleasant to listen to.

Another limitation was the use of a free-response question as the response for the language sample prompt. Each speaker answered the question differently and with their own personal prosody and inflections. A standardized passage for the language sample speakers to read would help to minimize the differences between speakers. In addition, both German speakers were graduate students in their mid-20s, while both Dutch speakers were in their early 50s. Using speakers of similar ages would also help to standardize the language samples.

Even better than trying to minimize differences in speakers as a group would be to use the same speaker for multiple languages in application of what is known as the matched-guise technique, wherein a single speaker with native-level language ability and accent in more than one language produce the samples for more than one language. This would wholly eliminate differences in speaker age, prosody, pitch, and several other factors. However, finding speakers with the necessary level of native proficiency and native accent in two languages is difficult, which is why this method was not employed in this study. Another option would be to use a

larger group of more closely matched speakers for a larger variety of voices that would theoretically average out any remaining differences.

A final limitation of this study was found in the gender of the speaker samples: the female German speaker had harsher and more strongly correlated ratings for the same "negative" stereotypes than the male speaker, suggesting that participants are judging the male and female speaker voices differently along gender lines as well, with traditionally "masculine" traits (Militaristic, Strict, Serious) not being as negatively correlated with the AR as in the female speaker samples.

Conclusion

Language attitude study considers not only the instantaneous state of the listener and the speaker in a conversation, but also their histories, experiences, and socialization to determine how people perceive and use language with speakers of diverse linguistic backgrounds. Drawing on this critical background, this research project explored whether the perception of German being “harsh” is due to cultural and stereotypical notions that may subconsciously but actively affect a listener’s perception of what could otherwise be a potentially neutral or at the very least less polarizing stimulus.

This was accomplished using a survey that asks respondents to score their agreement with stereotypes about Germans or Dutch people and rate language samples of German and Dutch aesthetically on a Likert scale. The data from these groups were compared against each other and analyzed using t-tests and Pearson product-moment correlations. The group primed with German stereotypes were hypothesized to rate whatever language follows more “harshly” and negatively aesthetically, regardless of the language’s real identity. This hypothesis was not supported by the data. Participants who rated the unlabeled language samples without stereotype prompting are hypothesized to rate the language less harshly and make more neutral stereotype judgements. The first part of this hypothesis was not supported by the data, but the second part was: the unlabeled male speaker sample of German elicited a significantly lower GSA mean than the labeled condition. Finally, the unlabeled languages were hypothesized to be rated more neutrally overall, refuting the inherent value hypothesis. This was also not shown to be the case, as there were no significant differences between the AR means across labeled, unlabeled, and switch conditions either intra- or cross-linguistically. Individual speaker differences made most

of the difference in the AR ratings, as the male Dutch voice had a significantly lower AR score and the female Dutch voice a significantly higher AR score than the other samples.

Although the stereotype of German as a "harsh" or unfriendly language will persist in American culture, there is no less reason to research the motivations and processes behind this notion. By systematically testing listener's reactions to both labeled and unlabeled language samples, it was hoped that this study would lead to a deeper understanding of the factors that influence our subconscious processing of language and the conscious attitudes formed from it. This is especially relevant not only in the field of Second Language Acquisition, where the perceptions of certain languages being "harsh" or "beautiful" could sway enrollment numbers or affect learner engagement, but more broadly in a world where globalization no longer the exception but the norm and where people must communicate more and more often across cultural and linguistic borders.

Addressing these stereotypes head-on pedagogically with sound sociolinguistic research could go a long way towards improving intercultural competency not just in world language classrooms, but across such diverse disciplines as history, sociology, psychology, and political science. Research has shown that although linguistically based in-group out-group biases are still very much alive and well, that with concerted effort, the notion that various linguistic and people groups are fundamentally foreign and dissimilar can be broken down through language study, cultural immersion and other avenues in order to sensitize learners of all ages to approaching cultural differences throughtfully and productively (Shin, Leal, & Ellison, 2015; Medina-Lopez-Portillo, 2004). An insight into how language attitudes are generated is not only timely but necessary if we are to best understand each other into the future.

References

Bradac, J. (1990). Language Attitudes and Impression Formation. *Handbook of Language and Social Psychology*, 387-412.

Cargile, A., Giles, H., Ryan, E., & Bradac, J. (1994). Language Attitudes as a Social Process: A Conceptual Model and New Directions. *Language and Communication, 14*(3), 211-236.

Copy Cat Channel. (2013, July 19). How German Sounds Compared To Other Languages. Retrieved from https://www.youtube.com/watch?v=-_xUIDRxdmc

Cronin, M. (2013, August 3). German Sounds Harsher Than Other Languages, And Here's Why. HuffPost. Retrieved from https://www.huffpost.com/entry/german-harsh-language-other-languages-video_n_3683379

Giles, H., Bourhis, R., & Davies, A. (1979). Prestige speech styles: the imposed norm and inherent value hypotheses. *Language and society: anthropological issues*, 589-596.

Giles, H., Bourhis, R., & Taylor, D. (1977). Towards a theory of language in ethnic group relations. In H. Giles, *Language, Ethnicity and Intergroup Relations* (pp. 307-344). Harcourt Brace Jovanovich.

Giles, H., Bourhis, R., Trudgill, P., & Lewis, A. (1974). The imposed norm hypothesis: a validation. *The Quarterly Journal of Speech, 60*(4), 405-410.

Halliday, M., McIntosh, A., & Strevens, P. (1964). *The Linguistic Sciences and Language Teaching.* London: Longman.

Jenkin, M. (2014, July 17). What makes a language attractive – its sound, national identity or familiarity? The Guardian. Retrieved from https://www.theguardian.com/education/2014/jul/17/what-makes-a-language-attractive

Kramer, R. (1992). Ethnolinguistic vitality perception in Inrael in the wake of the Intifada. *Journal of Multilingual and Multicultural Development*, 487-503.

Medina-Lopez-Portillo, A. (2004). Intercultural Learning Assessment: The Link between Program Duration and the Development of Intercultural Sensitivity. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 179-199.

rewboss. (2014, November 16). The harsh German language. Retrieved from https://www.youtube.com/watch?v=PuxgMGBymos

Ryan, E., Giles, H., & Hewstone, M. (1988). The Measurement of Language Attitudes. *Sociolinguistcs: an International Handbook of the Science of Language and Society*, 1068-1081.

Schulz, R. A., & Haerle, B. M. (1995). Beer, Fast Cars, and...: Stereotypes Held by U.S. College-Level Students of German. *Unterrichtspraxis*, 29-39.

Shin, H., Leal, D. L., & Ellison, C. (2015). Does Anti-Hispanic Bias Motivate Opposition to Non-English Languages? *Sociological Inquiry*, 375-406.

Trudgill, P., & Giles, H. (1978). Sociolinguistics and Linguistic Value Judgements: Correctness, Adequacy, and Aesthetics. *The Functions of Language and Literature Studies*.

Zahn, C., & Hopper, R. (1985). Measuring Language Attitudes: The Speech Evaluation Instrument. *Journal of Language and Social Psychology, 4*(2).

Zudeick, P. (2012, December 9). *Germans and Punctuality*. Retrieved from Deutsche Welle: https://www.dw.com/en/germans-and-punctuality/a-16430264

Zudeick, P. (2012, November 19). *Order Makes Germans' World Go Round*. Retrieved from Deutsche Welle: https://www.dw.com/en/order-makes-germans-world-go-round/a-16388879

Zudeick, P. (2012, December 17). *The Environmentally Friendly German*. Retrieved from Deutsche Welle: https://www.dw.com/en/the-environmentally-friendly-german/a-16459092

Zudeick, P. (2012, November 26). *Those Frugal Germans*. Retrieved from Deutsche Welle: https://www.dw.com/en/those-frugal-germans/a-16398388

Appendix

German as German, Group 1M (Individual Stereotype Correlations)								
Stereotype	Militaristic		Eco-Friendly		Strict		Serious	
Comparison Mean	GSA	AR	GSA	AR	GSA	AR	GSA	AR
Correlation Coefficent (r)	0.278	-0.019	0.418	0.256	0.686	0.105	0.544	-0.058
p-value	0.382	0.954	0.137	0.377	0.005	0.708	0.036	0.839
95% CI	-0.352	-0.586	-0.145	-0.318	0.268	-0.430	0.044	-0.554
	0.735	0.561	0.776	0.692	0.887	0.586	0.826	0.469
t-value	0.915	-0.060	1.593	0.917	3.399	0.382	2.336	-0.208
df	10		12		13		13	
Stereotype	Heavy Alcohol Consumers		Innovative		Intelligent		Greedy	
Comparison Mean	GSA	AR	GSA	AR	GSA	AR	GSA	AR
Correlation Coefficent (r)	0.497	0.141	0.418	0.605	0.191	0.293	0.198	-0.029
p-value	0.050	0.603	0.155	0.028	0.479	0.270	0.498	0.922
95% CI	0.002	-0.381	-0.173	0.081	-0.337	-0.237	-0.372	-0.551
	0.797	0.595	0.788	0.867	0.627	0.689	0.659	0.510
t-value	2.144	0.532	1.526	2.520	0.727	1.148	0.698	-0.100
df	14		11		14		12	
Stereotype	Egalitarian		Logical		Organized		Rich	
Comparison Mean	GSA	AR	GSA	AR	GSA	AR	GSA	AR
Correlation Coefficent (r)	-0.355	0.041	0.333	-0.238	0.817	0.614	0.371	0.151
p-value	0.213	0.888	0.266	0.433	1.11E-04	0.011	0.157	0.577
95% CI	-0.745	-0.500	-0.267	-0.698	0.541	0.170	-0.153	-0.373
	0.216	0.560	0.747	0.360	0.934	0.851	0.732	0.602
t-value	-1.316	0.144	1.172	-0.814	5.308	2.912	1.495	0.571
df	12		11		14		14	
Stereotype	Cultured		Punctual					
Comparison Mean	GSA	AR	GSA	AR				
Correlation Coefficent (r)	0.465	0.198	0.745	0.787				
p-value	0.060	0.446	0.009	0.004				
95% CI	-0.020	-0.312	0.262	0.354				
	0.773	0.620	0.929	0.942				
t-value	2.037	0.783	3.347	3.823				
df	15		9					

German as German, Group 1F (Individual Stereotype Correlations)								
Stereotype	Militaristic		Eco-Friendly		Strict		Serious	
Comparison Mean	GSA	AR	GSA	AR	GSA	AR	GSA	AR
Correlation Coefficent (r)	0.100	-0.597	0.555	0.652	0.515	-0.542	0.620	-0.443
p-value	0.722	0.019	0.096	0.041	0.050	0.037	0.010	0.086
95% CI	-0.434	-0.850	-0.115	0.038	0.004	-0.825	0.179	-0.770
	0.583	-0.123	0.878	0.909	0.813	-0.041	0.853	0.068
t-value	0.364	-2.686	1.885	2.432	2.166	-2.323	2.955	-1.848
df	13		8		13		14	
Stereotype	Heavy Alcohol Consumers		Innovative		Intelligent		Greedy	
Comparison Mean	GSA	AR	GSA	AR	GSA	AR	GSA	AR
Correlation Coefficent (r)	0.671	-0.480	0.346	0.444	0.056	0.571	0.711	-0.552
p-value	0.006	0.070	0.226	0.112	0.849	0.033	0.006	0.050
95% CI	0.242	-0.796	-0.226	-0.113	-0.489	0.058	0.264	-0.846
	0.881	0.043	0.741	0.789	0.570	0.845	0.907	-0.002
t-value	3.263	-1.973	1.278	1.716	0.195	2.409	3.357	-2.197
df	13		12		12		11	
Stereotype	Egalitarian		Logical		Organized		Rich	
Comparison Mean	GSA	AR	GSA	AR	GSA	AR	GSA	AR
Correlation Coefficent (r)	-0.239	0.569	0.544	-0.382	0.406	-0.051	0.583	0.280
p-value	0.411	0.034	0.029	0.144	1.50E-01	0.862	0.029	0.332
95% CI	-0.683	0.055	0.067	-0.738	-0.159	-0.566	0.076	-0.294
	0.334	0.844	0.819	0.140	0.770	0.493	0.850	0.706
t-value	-0.852	2.395	2.428	-1.547	1.538	-0.178	2.484	1.011
df	12		14		12		12	
Stereotype	Cultured		Punctual					
Comparison Mean	GSA	AR	GSA	AR				
Correlation Coefficent (r)	-0.067	0.425	0.683	0.008				
p-value	0.812	0.114	0.014	0.981				
95% CI	-0.560	-0.112	0.180	-0.569				
	0.461	0.770	0.903	0.579				
t-value	-0.243	1.693	2.957	0.024				
df	13		10					

German as German, Group 1 Overall (Individual Stereotype Correlations)								
Stereotype	Militaristic		Eco-Friendly		Strict		Serious	
Comparison Mean	GSA	AR	GSA	AR	GSA	AR	GSA	AR
Correlation Coefficent (r)	0.174	-0.397	0.428	0.445	0.589	-0.216	0.495	-0.202
p-value	0.384	0.040	0.037	0.030	0.001	0.252	0.005	0.275
95% CI	-0.220	-0.675	0.030	0.050	0.291	-0.534	0.170	-0.519
	0.520	-0.020	0.709	0.719	0.783	0.157	0.723	0.164
t-value	0.886	-2.163	2.222	2.328	3.860	-1.169	3.067	-1.113
df	25		22		28		29	
Stereotype	Heavy Alcohol Consumers		Innovative		Intelligent		Greedy	
Comparison Mean	GSA	AR	GSA	AR	GSA	AR	GSA	AR
Correlation Coefficent (r)	0.531	-0.122	0.368	0.503	0.152	0.437	0.431	-0.326
p-value	0.002	0.512	0.059	0.008	0.424	0.016	0.025	0.097
95% CI	0.218	-0.457	-0.013	0.152	-0.221	0.092	0.060	-0.628
	0.745	0.243	0.657	0.741	0.485	0.689	0.697	0.062
t-value	3.378	-0.664	1.982	2.908	0.812	2.574	2.385	-1.724
df	29		25		28		25	
Stereotype	Egalitarian		Logical		Organized		Rich	
Comparison Mean	GSA	AR	GSA	AR	GSA	AR	GSA	AR
Correlation Coefficent (r)	-0.279	0.373	0.451	-0.302	0.717	0.301	0.411	0.207
p-value	0.151	0.051	0.014	0.111	8.31E-06	0.106	0.024	0.273
95% CI	-0.590	0.000	0.101	-0.602	0.481	-0.066	0.060	-0.166
	0.105	0.655	0.701	0.072	0.856	0.597	0.672	0.528
t-value	-1.480	2.048	2.623	-1.649	5.441	1.671	2.387	1.118
df	26		27		28		28	
Stereotype	Cultured		Punctual					
Comparison Mean	GSA	AR	GSA	AR				
Correlation Coefficent (r)	0.199	0.336	0.710	0.361				
p-value	0.275	0.060	0.000	0.090				
95% CI	-0.161	-0.014	0.420	-0.060				
	0.512	0.613	0.868	0.673				
t-value	1.113	1.956	4.616	1.775				
df	30		21					

Groups 1 and 2 Overall (Individual Stereotype vs. GSA Correlations)- German							
Stereotype	Militaristic	Eco-Friendly	Strict	Serious	Heavy Alcohol	Innovative	Intelligent
Correlation Coefficent (r)	0.200	0.321	0.447	0.434	0.547	-0.012	0.295
p-value	0.155	0.028	0.001	0.001	1.30E-05	0.933	0.027
95% CI	-0.077	0.037	0.200	0.198	0.331	-0.279	0.035
	0.448	0.557	0.640	0.622	0.708	0.257	0.518
t-value	1.443	2.272	3.564	3.605	4.798	-0.084	2.268
df	50	45	51	56	54	52	54
Stereotype	Greedy	Egalitarian	Logical	Organized	Rich	Cultured	Punctual
Correlation Coefficent (r)	0.238	-0.253	0.477	0.642	0.523	0.311	0.670
p-value	0.107	0.076	0.000	2.17E-07	4.16E-05	0.017	1.70E-06
95% CI	-0.053	-0.497	0.237	0.450	0.299	0.060	0.456
	0.492	0.027	0.662	0.778	0.692	0.525	0.810
t-value	1.645	-1.814	3.873	5.985	4.469	2.470	5.629
df	45	48	51	51	53	57	39

Groups 1 and 2 Overall (Individual Stereotype vs. GSA Correlations)- Dutch							
Stereotype	Militaristic	Eco-Friendly	Strict	Serious	Heavy Alcohol Consum	Innovative	Intelligent
Correlation Coefficent (r)	0.198	0.267	0.393	0.370	0.011	0.506	0.609
p-value	0.148	0.061	0.006	0.011	0.945	0.00032874	2.14E-06
95% CI	-0.071	-0.012	0.119	0.092	-0.291	0.253	0.400
	0.440	0.508	0.611	0.594	0.310	0.695	0.757
t-value	1.468	1.918	2.866	2.670	0.069	3.896	5.371
df	53	48	45	45	41	44	49
Stereotype	Greedy	Egalitarian	Logical	Organized	Rich	Cultured	Punctual
Correlation Coefficent (r)	0.202	0.033	0.483	0.583	0.608	0.334	0.585
p-value	0.188	0.831	0.00089621	2.10E-05	3.64E-06	0.019	0.001
95% CI	-0.101	-0.263	0.217	0.353	0.394	0.058	0.291
	0.471	0.323	0.682	0.747	0.759	0.562	0.778
t-value	1.339	0.214	3.575	4.763	5.247	2.430	3.886
df	42	43	42	44	47	47	29

German Unlabeled, Group 3 Overall (Individual Stereotype Correlations)								
Stereotype	Militaristic		Eco-Friendly		Strict		Serious	
Comparison Mean	GSA	AR	GSA	AR	GSA	AR	GSA	AR
Correlation Coefficent (r)	-0.232	-0.467	0.457	-0.034	0.558	-0.157	0.476	-0.116
p-value	0.254	0.016	0.025	0.876	0.002	0.417	0.009	0.547
95% CI	-0.568	-0.723	0.065	-0.431	0.240	-0.495	0.133	-0.463
	0.171	-0.097	0.726	0.375	0.767	0.223	0.718	0.261
t-value	-1.168	-2.586	2.408	-0.158	3.492	-0.824	2.814	-0.609
df	24		22		27		27	
Stereotype	Heavy Alcohol Consumers		Innovative		Intelligent		Greedy	
Comparison Mean	GSA	AR	GSA	AR	GSA	AR	GSA	AR
Correlation Coefficent (r)	0.623	-0.526	0.318	0.161	0.523	0.561	0.062	-0.467
p-value	0.002	0.012	0.121	0.442	0.004	0.002	0.772	0.021
95% CI	0.273	-0.775	-0.088	-0.250	0.193	0.246	-0.350	-0.732
	0.827	-0.134	0.634	0.523	0.746	0.770	0.454	-0.078
t-value	3.561	-2.764	1.609	0.782	3.187	3.526	0.293	-2.477
df	20		23		27		22	
Stereotype	Egalitarian		Logical		Organized		Rich	
Comparison Mean	GSA	AR	GSA	AR	GSA	AR	GSA	AR
Correlation Coefficent (r)	0.306	0.138	0.247	-0.028	0.319	0.227	0.484	0.050
p-value	0.178	0.552	0.197	0.885	0.091	0.237	0.019	0.821
95% CI	-0.145	-0.313	-0.131	-0.391	-0.053	-0.153	0.089	-0.370
	0.652	0.537	0.563	0.342	0.614	0.548	0.747	0.453
t-value	1.400	0.605	1.324	-0.146	1.751	1.209	2.532	0.229
df	19		27		27		21	
Stereotype	Cultured		Punctual					
Comparison Mean	GSA	AR	GSA	AR				
Correlation Coefficent (r)	0.402	0.466	0.753	0.089				
p-value	0.038	0.014	3.366E-05	0.687				
95% CI	0.026	0.104	0.494	-0.336				
	0.678	0.719	0.889	0.483				
t-value	2.196	2.632	5.245	0.409				
df	25		21					

Acknowledgements

First and foremost, a huge thank you to my fantastic advisor, Prof. Carmen Taleghani-Nikazm, for your endless encouragement, mentorship, and guidance both over the course of this project and for my entire undergraduate career. Your incredibly specific and helpful feedback at every stage of the process is what really took this thesis to the next level and made it one of the most important learning experiences I've ever had. I would not have been able to get this project off the ground without you.

Next, I'd like to acknowledge and thank Professors Anna Grotans and Terry Gustafson for serving on my Defense Committee and for being there for me here at Ohio State from Day 1. As my current and ex-major advisors, you've been very important figures in my experience here through all the twists and turns my majors have thrown at me.

I would also like to thank the entire Department of Germanic Languages and Literatures for repeatedly letting me drop by your offices unannounced to ramble, say hello, and bounce ideas around. Thank you for sharing in my excitement for this project and for being a fantastic academic home the past four years!

I am also grateful to several people in the Department of Linguistics: Dr. Kathryn Campbell-Kibler for her advising on the more technical linguistic aspects of this study, Dr. Cynthia Clopper for the use of her lab, and to Dr. Hope Dawson for helping me set up the LOC pool. A special thank you also goes to Josephine Templer, Johannes Vith, Dr. Mark Moritz, Dr. Bernadette Vankeerbergen, and Merijn Vanderheijden, who volunteered their time to record themselves for my language samples.

Un grandissimo grazie mille, Andrea, non solo per la tua pazienza e le lezioni d'italiano, ma anche per trascorrere tanti (troppi) weekend con me su Skype per risolvere problemi di R e per aiutarmi con l'analisi dei dati. Tu sei un amico incredibile e sono molto grata che la nostra amicizia sia sopravvissuta alle molte, molte linee di quel maledetto codice scritto per creare un metodo di analisi funzionante (e che non mi facesse impazzire!). Ringrazio anche Erika, la mia italiana preferita: mi hai supportato immensamente con i tuoi "heeellooo!" nonostante fossi impegnata in quel momento a scrivere la tua tesi. Non dimenticherò mai la tua amicizia salda e i tuoi sorrisi smaglianti in questo periodo speciale della mia vita. Ce l'abbiamo fatta![i]

Finally, perhaps the most important thank you goes to my parents, who encouraged my love of learning and endless parade of diverse and obscure interests from the day I was born. German isn't what I ever thought I would settle on, but without your unflinching and deeply loving support, I wouldn't have been able to pursue my passions as fearlessly and wholeheartedly in the way that has made not only this thesis, but my future livelihood possible. You instilled me early a sense of independence, responsibility, and ambition that has and will continue serve me well throughout my life. Your trust in me to make these difficult and important decisions for my future is one of the most valuable gifts you could have ever given me. Thank you, so very, very much for being the kind of role models and cheerleaders any daughter could ever dream of. Thank you.

[i] Thank you so incredibly much, Andrea, not only for your patience and the Italian lessons, but also for spending too many of your weekends with me on Skype troubleshooting R and helping me with my data analysis. You're an incredible friend and I'm so grateful our friendship survived the many, many lines of botched code in the process of creating an analysis method that worked (and prevented me from going crazy!). I'd also like to thank Erika, my favorite Italian: you supported me immensely with your "heelllooos!" despite working on your own thesis at the same time. I will never forget your steadfast friendship and goofy smiles in this special time in my life. We did it!